Barriers and Local Minima in Energy Landscapes

Petteri Kaski

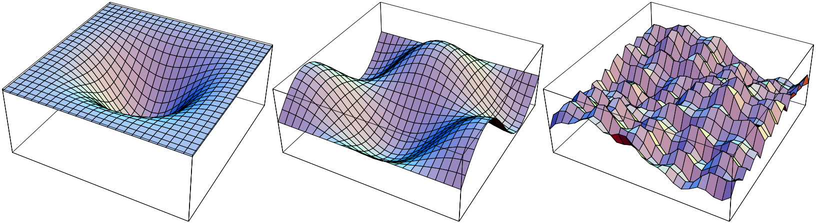
Helsinki Institute for Information Technology HIIT
University of Helsinki, Department of Computer Science
petteri.kaski@cs.helsinki.fi

Mariehamn, Tue July 17, 2007

Contents

1. Constraint satisfaction problems (K-SAT)

2. Energy landscape of a K-SAT instance



3. Analysis of barriers in random K-SAT landscapes ($K = 8$)

   Inspiration/Techniques: Proof of clustering of solutions in random K-SAT by Mézard, Mora, and Zecchina (2005)

Constraint satisfaction problems (CSPs)

- Task: Given M constraints over N variables $x_1, \ldots, x_N$, assign values to variables so that all constraints are satisfied
- **K-satisfiability** (**K-SAT**):
  - all variables take values in $\{0, 1\}$
  - each constraint (*clause*) forbids one combination of values to some K variables
- **Example** ($K = 3$, $M = 16$, $N = 4$):

$$(\bar{x}_1 \vee x_2 \vee x_3)(x_1 \vee \bar{x}_2 \vee x_3)(x_1 \vee x_2 \vee \bar{x}_3)(\bar{x}_1 \vee \bar{x}_2 \vee \bar{x}_3)$$
$$(\bar{x}_1 \vee x_2 \vee x_4)(x_1 \vee \bar{x}_2 \vee x_4)(x_1 \vee x_2 \vee \bar{x}_4)(\bar{x}_1 \vee \bar{x}_2 \vee \bar{x}_4)$$
$$(\bar{x}_1 \vee x_3 \vee x_4)(x_1 \vee \bar{x}_3 \vee x_4)(x_1 \vee x_3 \vee \bar{x}_4)(\bar{x}_1 \vee \bar{x}_3 \vee \bar{x}_4)$$
$$(\bar{x}_2 \vee x_3 \vee x_4)(x_2 \vee \bar{x}_3 \vee x_4)(x_2 \vee x_3 \vee \bar{x}_4)(\bar{x}_2 \vee \bar{x}_3 \vee \bar{x}_4)$$

- $(x_1 \vee \bar{x}_2 \vee x_4)$ forbids the combination $x_1 = 0$, $x_2 = 1$, $x_4 = 0$

Landscape of a K-SAT instance

- Assume an instance of K-SAT

$$(\bar{x}_1 \vee x_2 \vee x_3)(x_1 \vee \bar{x}_2 \vee x_3)(x_1 \vee x_2 \vee \bar{x}_3)(\bar{x}_1 \vee \bar{x}_2 \vee \bar{x}_3)$$
$$(\bar{x}_1 \vee x_2 \vee x_4)(x_1 \vee \bar{x}_2 \vee x_4)(x_1 \vee x_2 \vee \bar{x}_4)(\bar{x}_1 \vee \bar{x}_2 \vee \bar{x}_4)$$
$$(\bar{x}_1 \vee x_3 \vee x_4)(x_1 \vee \bar{x}_3 \vee x_4)(x_1 \vee x_3 \vee \bar{x}_4)(\bar{x}_1 \vee \bar{x}_3 \vee \bar{x}_4)$$
$$(\bar{x}_2 \vee x_3 \vee x_4)(x_2 \vee \bar{x}_3 \vee x_4)(x_2 \vee x_3 \vee \bar{x}_4)(\bar{x}_2 \vee \bar{x}_3 \vee \bar{x}_4)$$

- A **state** s is an assignment of values to variables
- The **energy** $E(s)$ is the number of constraints violated by s

$$E(0100) = 3$$

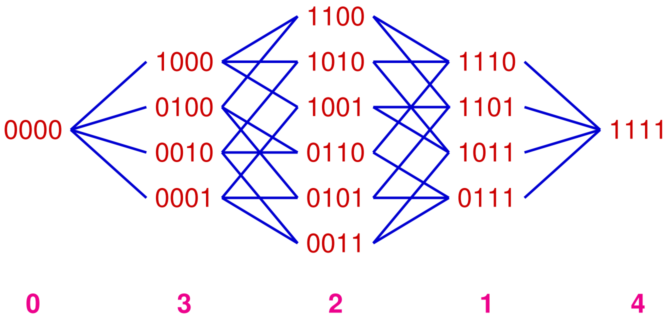
- Two states are **adjacent** if they differ in the value of exactly one variable

An example K-SAT landscape

- Instance ($K = 3$, $M = 16$, $N = 4$):

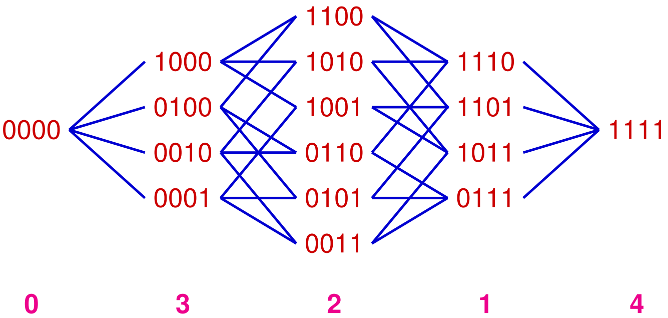
$$(\bar{x}_1 \vee x_2 \vee x_3)(x_1 \vee \bar{x}_2 \vee x_3)(x_1 \vee x_2 \vee \bar{x}_3)(\bar{x}_1 \vee \bar{x}_2 \vee \bar{x}_3)$$
$$(\bar{x}_1 \vee x_2 \vee x_4)(x_1 \vee \bar{x}_2 \vee x_4)(x_1 \vee x_2 \vee \bar{x}_4)(\bar{x}_1 \vee \bar{x}_2 \vee \bar{x}_4)$$
$$(\bar{x}_1 \vee x_3 \vee x_4)(x_1 \vee \bar{x}_3 \vee x_4)(x_1 \vee x_3 \vee \bar{x}_4)(\bar{x}_1 \vee \bar{x}_3 \vee \bar{x}_4)$$
$$(\bar{x}_2 \vee x_3 \vee x_4)(x_2 \vee \bar{x}_3 \vee x_4)(x_2 \vee x_3 \vee \bar{x}_4)(\bar{x}_2 \vee \bar{x}_3 \vee \bar{x}_4)$$

- States, adjacency, and energy:



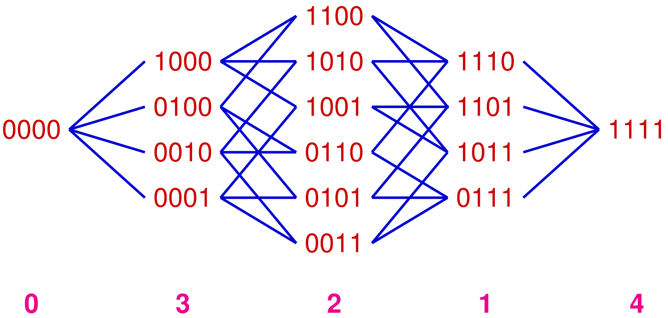
Energy barriers

- Let *s* and *t* be states
- Walking from *s* to *t* (via consecutive adjacent states) may require traversing states with higher energy



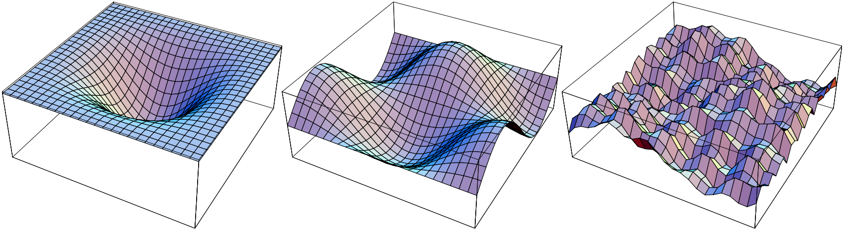
Energy barrier

- The **energy barrier** $B(s,t)$ separating s and t is the minimum increase in energy over $\max(E(s),E(t))$ required by *any walk* from s to t
- $B(1110,0000) = B(0000,1110) = 2$



Question:

Given a random instance of K-SAT, what does the landscape "look like" w.h.p. as $N \to \infty$?



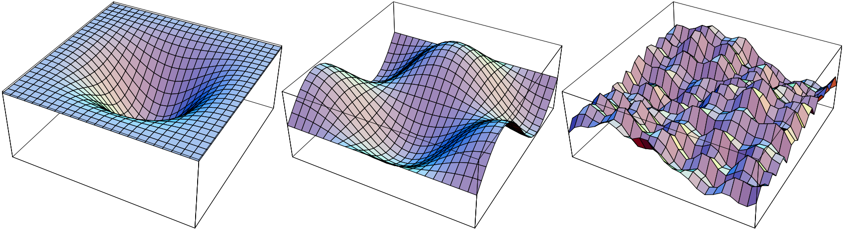
Random K-SAT

- Parameters: K, M, and N
- A random instance of K-SAT:

  Select M clauses independently and uniformly at random (with replacement) from the set of all possible clauses of length K over the N variables
- Clause density $\alpha = M/N$
- Keep K and α fixed, let $N \to \infty$

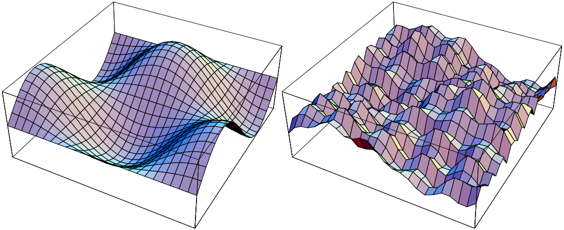
Question:

How does $\max_{s,t} B(s,t)$ scale as $N \to \infty$?



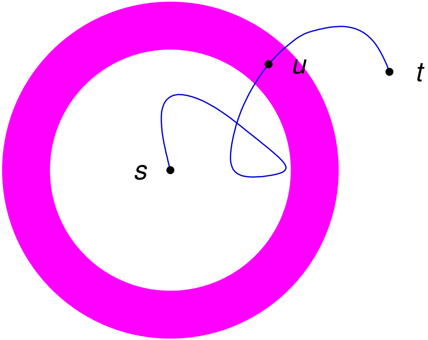
A first (?) rigorous lower bound:

For $K = 8$ there exists an interval in $\alpha \leq \alpha_{\mathrm{sat}}(K)$ such that w.h.p. a random K-SAT landscape has $\max_{s,t} B(s,t) = \Omega(N)$



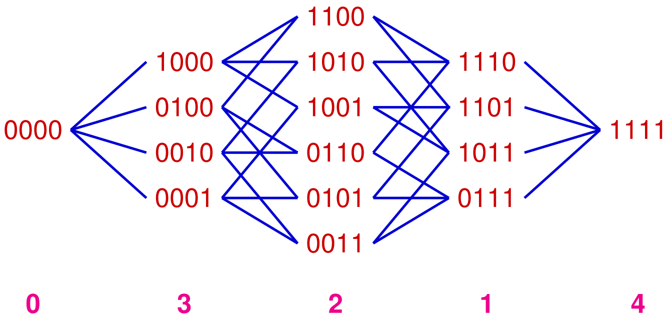
Proof strategy

- Assume s and t are two states with $E(s) = E(t) = 0$
- Assume s and t are "far away" from each other
- Prove that w.h.p. *every state* u at "intermediate distance" from s has $E(u) \gg 0$



Distance

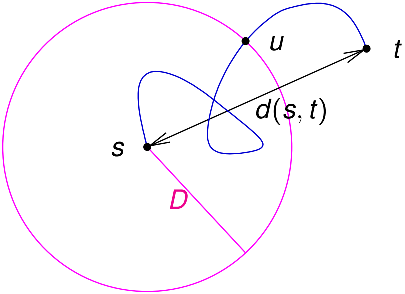
- The (Hamming) **distance** $d(s,t)$ is the number of variables in which s and t differ
- $d(1000,1110) = 2$



Proof strategy in more detail

- Observation:

  Every walk from s to t has *at least one state* u at each distance $D = 0, 1, \ldots, d(s, t)$ from s



- Prove that for some D and $\tau > 0$,
  w.h.p. *every state* u with $d(s, u) = D$ has $E(u) \geq \tau N$

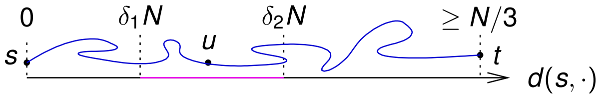
In yet more precise terms ...

- Mézard, Mora, and Zecchina (2005):

  For $K = 8$ and $\alpha \leq 170$,
  a random K-SAT landscape has w.h.p. a pair (s, t)
  with $E(s) = E(t) = 0$ and $d(s, t) \geq N/3$

- It thus suffices to show:

  For $K = 8$ and $\alpha \geq 167$,
  there exist $0 < \delta_1 < \delta_2 < 1/3$ and $\tau > 0$ such that
  a random K-SAT landscape has w.h.p. **no pair** (s, u)
  with $E(s) = 0$, $E(u) \leq \tau N$, and $\delta_1 N \leq d(s, u) \leq \delta_2 N$



Proof: Roadmap

- We follow footsteps of Mézard, Mora, and Zecchina (2005)
- Steps:
  1. the 1st moment method
     ($\Pr(X \geq 1) \leq \mathrm{Ex}(X)$ if X is a nonnegative rv),
  2. linearity of expectation,
  3. symmetry,
  4. independence, and
  5. counting

Step 1: The 1st moment method

- Let $Z_{K,M,N}$ be a random K-SAT instance
- Call a pair (s, u) of states (D, T)-**bad** for $Z_{K,M,N}$ if
  1. $d(s, u) = D$;
  2. $E(s) = 0$; and
  3. $E(u) \leq T$
- Let $b(Z_{K,M,N}, D, T)$ be the number of (D, T)-bad pairs for $Z_{K,M,N}$
- The 1st moment method:

$$\Pr\big(b(Z_{K,M,N}, D, T) \geq 1\big) \leq \mathrm{Ex}\big(b(Z_{K,M,N}, D, T)\big)$$

Step 2: Linearity of expectation

- Write $b(Z_{K,M,N}, D, T)$ as a sum of 0/1 indicators, one for each candidate pair (s, u):

$$b(Z_{K,M,N}, D, T) = \sum_{d(s,u)=D} \big[E(s) = 0 \;\&\; E(u) \leq T\big]$$

- Linearity of expectation:

$$\begin{aligned}
\mathrm{Ex}\big(b(Z_{K,M,N}, D, T)\big) &= \sum_{d(s,u)=D} \mathrm{Ex}\big(\big[E(s) = 0 \;\&\; E(u) \leq T\big]\big) \\
&= \sum_{d(s,u)=D} \Pr\big(\big[E(s) = 0 \;\&\; E(u) \leq T\big]\big)
\end{aligned}$$

Step 3: Symmetry

- There are $2^N\binom{N}{D}$ pairs (s, u) with $d(s, u) = D$
- Clauses are selected uniformly at random
- By symmetry, the probability

$$p(K, M, N, D, T) = \Pr\big([E(s) = 0 \ \& \ E(u) \leq T]\big)$$

is the same regardless of the choice (s, u) with $d(s, u) = D$

- Thus,

$$\begin{aligned}\mathrm{Ex}\big(b(Z_{K,M,N}, D, T)\big) &= \sum_{d(s,u)=D} \Pr\big([E(s) = 0 \ \& \ E(u) \leq T]\big) \\ &= 2^N\tbinom{N}{D} p(K, M, N, D, T)\end{aligned}$$

Step 4: Independence

- We derive an upper bound for

$$p(K, M, N, D, T) = \Pr\big(\big[E(s) = 0 \ \&\ E(u) \le T\big]\big)$$

- For $E(s) = 0 \ \&\ E(u) \le T$ to occur,
  some $M - T$ clauses *must* be satisfied by both s and u

- By independence (and union bound)

$$p(K, M, N, D, T) \le \binom{M}{M-T} q(K, N, D)^{M-T}$$

  where

$$q(K, N, D) = \Pr\big(s \text{ and } u \text{ satisfy a random clause}\big)$$

Step 5: Counting

- We want an expression for

$$q(K,N,D) = \Pr(0^N \text{ and } 1^D 0^{N-D} \text{ satisfy a random clause})$$

- Counting gives

$$\begin{aligned} q(K,N,D) &= \frac{2^K\binom{N}{K} - \binom{N-D}{K} - 2\sum_{j\geq 1}\binom{D}{j}\binom{N-D}{K-j}}{2^K\binom{N}{K}} \\ &= 1 - 2^{1-K} + 2^{-K}\binom{N-D}{K}\binom{N}{K}^{-1} \\ &\leq q(K, D/N) \end{aligned}$$

where

$$q(K,\delta) = 1 - 2^{1-K} + 2^{-K}(1-\delta)^K$$

Wrapping up & Asymptotics (1 / 5)

- We now have the upper bound

$$\Pr\bigl(b(Z_{K,M,N}, D, T) \geq 1\bigr) \leq F(K, M, N, D, T)$$

where

$$F(K, M, N, D, T) = 2^N \binom{N}{D} \binom{M}{M-T} q\bigl(K, D/N\bigr)^{M-T}$$

and

$$q(K, \delta) = 1 - 2^{1-K} + 2^{-K}(1-\delta)^K$$

- Reparametrize: $M = \alpha N$, $D = \delta N$, $T = \tau N$

Wrapping up & Asymptotics (2 / 5)

- We obtain

$$\Pr\big(b(Z_{K,\alpha,N},\delta,\tau) \geq 1\big) \leq F(K,\alpha,N,\delta,\tau)$$

where

$$F(K,\alpha,N,\delta,\tau) = 2^N \binom{N}{\delta N}\binom{\alpha N}{\tau N} q(K,\delta)^{(\alpha-\tau)N}$$

and

$$q(K,\delta) = 1 - 2^{1-K} + 2^{-K}(1-\delta)^K$$

- Now let

$$f(K,\alpha,N,\delta,\tau) = \frac{1}{N}\log F(K,\alpha,N,\delta,\tau)$$

Wrapping up & Asymptotics (3 / 5)

- We have

$$f(K,\alpha,N,\delta,\tau) \le \log 2 + H(\delta) + \alpha H(\tau/\alpha) + \\ + (\alpha-\tau)\log q(K,\delta)$$

with

$$q(K,\delta) = 1 - 2^{1-K} + 2^{-K}(1-\delta)^K$$

and

$$H(\lambda) = -\lambda\log\lambda - (1-\lambda)\log(1-\lambda)$$

- A “small” fixed $\tau > 0$ gives $f(K,\alpha,N,\delta,\tau) \ll 0$ as $N \to \infty$, *uniformly* for all “large” α and all $\delta \in [\delta_1,\delta_2]$ for some fixed $0 < \delta_1 < \delta_2 < 1/3$

Wrapping up & Asymptotics (4 / 5)

- Thus, for a fixed K and a fixed $\tau > 0$ we have

$$\Pr\big(b(Z_{K,\alpha,N}, D, \tau N) \geq 1\big) = \exp\big(-\Omega(N)\big)$$

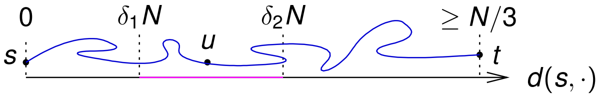
uniformly for all large enough α and all $\delta_1 N \leq D \leq \delta_2 N$

- Using the union bound (over $O(N)$ choices of D),

$$\Pr\left(\sum_{\delta_1 N \leq D \leq \delta_2 N} b(Z_{K,\alpha,N}, D, \tau N) \geq 1\right) = \exp\big(-\Omega(N)\big)$$

Wrapping up & Asymptotics (5 / 5)

- For $K = 8$ we can take $\tau = 10^{-6}$, any $\alpha \geq 167$, $\delta_1 = 0.1$, and $\delta_2 = 0.2$



- Mézard, Mora, and Zecchina (2005) show that for $K = 8$ and any $\alpha \leq 170$ a random K-SAT landscape has w.h.p. a pair (s, t) with $E(s) = E(t) = 0$ and $d(s, t) \geq N/3$
- For $K = 8$ and $167 \leq \alpha \leq 170$ a random K-SAT landscape has w.h.p. $\max_{s,t} B(s, t) \geq \tau N$ $\square$

Conclusion

- Extensive barriers separate distant solutions in random K-SAT for $K \geq 8$ and "large" $\alpha \leq \alpha_{\mathrm{sat}}(K)$
- Q: What happens for $K < 8$?
- Q: Do there exist states with (extensive) barriers separating them from *every* solution ?